Combining Omics and Cognitive Tests Data from SMC individuals, Artificial Intelligence Technology Identifies Genomic Biomarkers for Early Detection of Alzheimer's Disease.

Martin Kindermans, MSc[1], **Mohammad Afshar, MD, PhD[1]**, Patrizia Andrea Chiesa, PhD[2,3,4], Adrien Etcheto, MSc[1], Simone Lista, PhD[2,3,4], Pablo Lemercier[2,3,4], Frédéric Parmentier, PhD[1], Andrea Vergallo, MD[2,3,4], Coralie Williams, MSc[1] and Harald Hampel, Prof., MD, PhD[3]

(1) Ariana Pharma, Paris, France
(2) Institute of Memory and Alzheimer's Disease (IM2A), Department of Neurology, Pitié-Salpêtrière Hospital, AP-HO, Boulevard de l'hôpital, F-75013, Paris, France
(3) Sorbonne University, GRC n° 21, Alzheimer Precision Medicine (APM), AP-HP, Pitié-Salpêtrière Hospital, Boulevard de l'hôpital, F-75013, Paris, France
(4) Brain & Spine Institute (ICM), INSERM U 1127, CNRS UMR 7225, Boulevard de l'hôpital, F-75013, Paris, France

AAIC 2019, July 14-18
Los Angeles, CA, U.S.

Conflict of Interest Disclosure

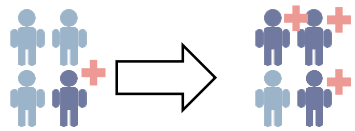**Mohammad Afshar** is employee and shareholder of Ariana Pharma

Introduction

Number of AD cases will triple, healthcare cost will double by 2030

2018   2030

Source: World Alzheimer Report 2018

Early detection and management of some, if not all, future AD cases is critical

Focusing on profiles at risk for AD, but not yet affected, will help us

Today!

The combination of DNA analysis and Artificial Intelligence is needed to analyze extensively characterized cohorts

We need a cohort of subjects, <u>not having AD</u>, MCI, or other neurological/psychiatric disorder, <u>but at risk</u>, with extensive DNA information
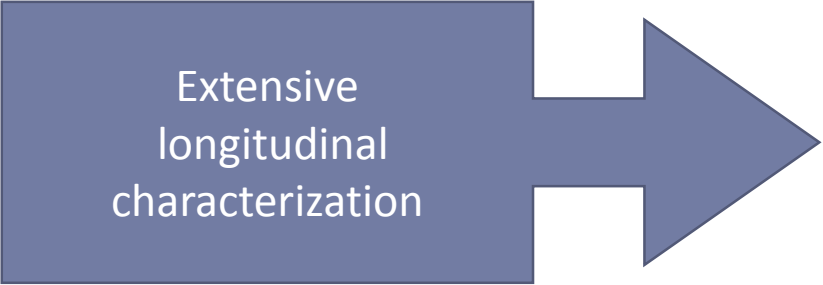
Introduction

Study performed in a large-scale, university-based, monocentric cohort

Cognitively and physically normal Caucasian individuals with SMC
SMC defined as positive response to both questions:
- Are you complaining about your memory?
- Is it a regular complaint which lasts more than 6 months?

☑ 318 subjects

☑ 70–85 years

☑ Subjective memory complaints

☑ Unimpaired cognition and memory

Extensive longitudinal characterization

Amyloid (18F-Florbetapir) PET Imaging

Brain MRI / fMRI imaging

Brain glucose-metabolism
[18F-fluorodeoxyglucose (18F-FDG) PET]

Cognitive assessment

PET: positron emission tomography ; MRI: magnetic resonance imaging; AD: Alzheimer's disease

Questions

Cognitive &
Memory tests

Neuroimaging

Variants

Blood-based
biomarkers
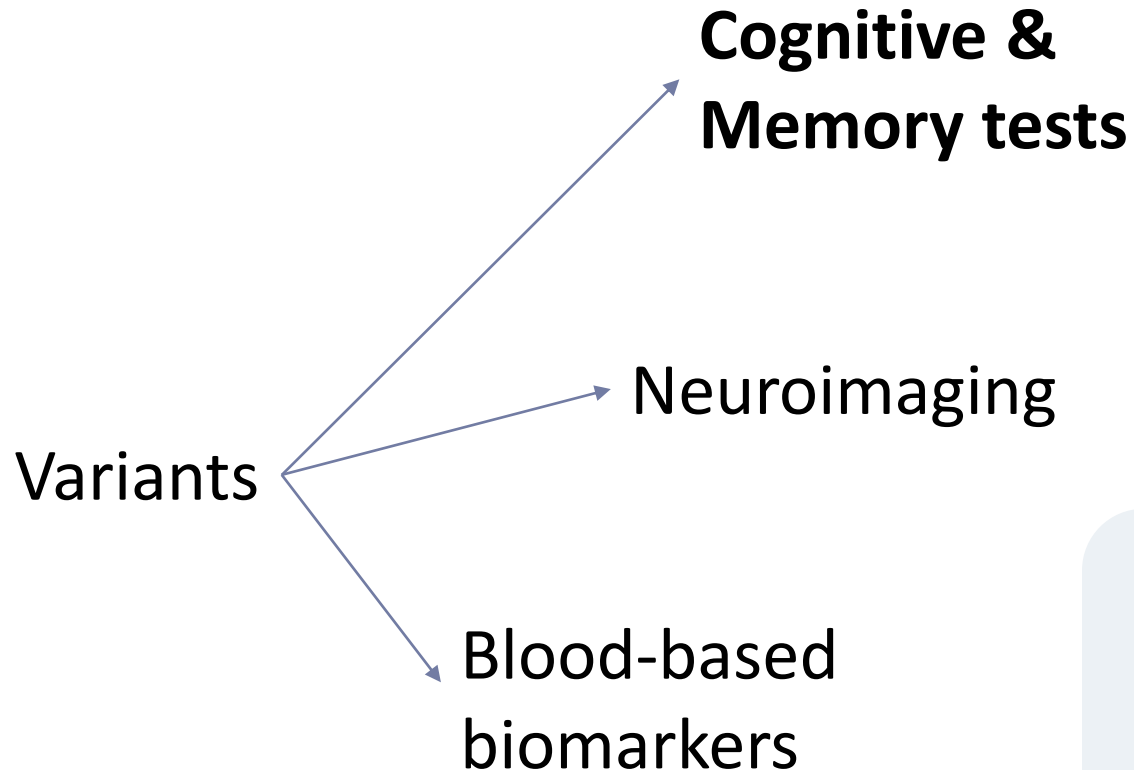
Hibar *et al*, Nature 2015

nature
International journal of science

Letter | Published: 21 January 2015

Common genetic variants influence
human subcortical brain structures

Can we identify genomic variants that are simultaneously linked to multiple types of endpoints?

Methodology

Systematic omic Analysis of longitudinal cohorts

~100 000s descriptors

~ 100s samples

Identification of biomarkers of response from clinical studies
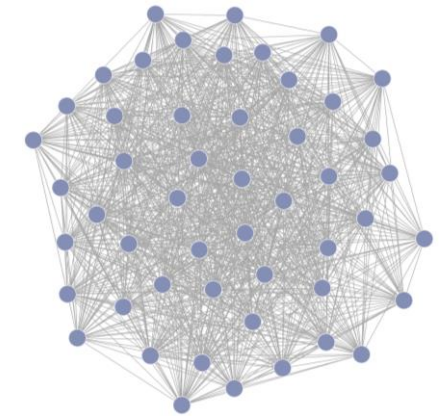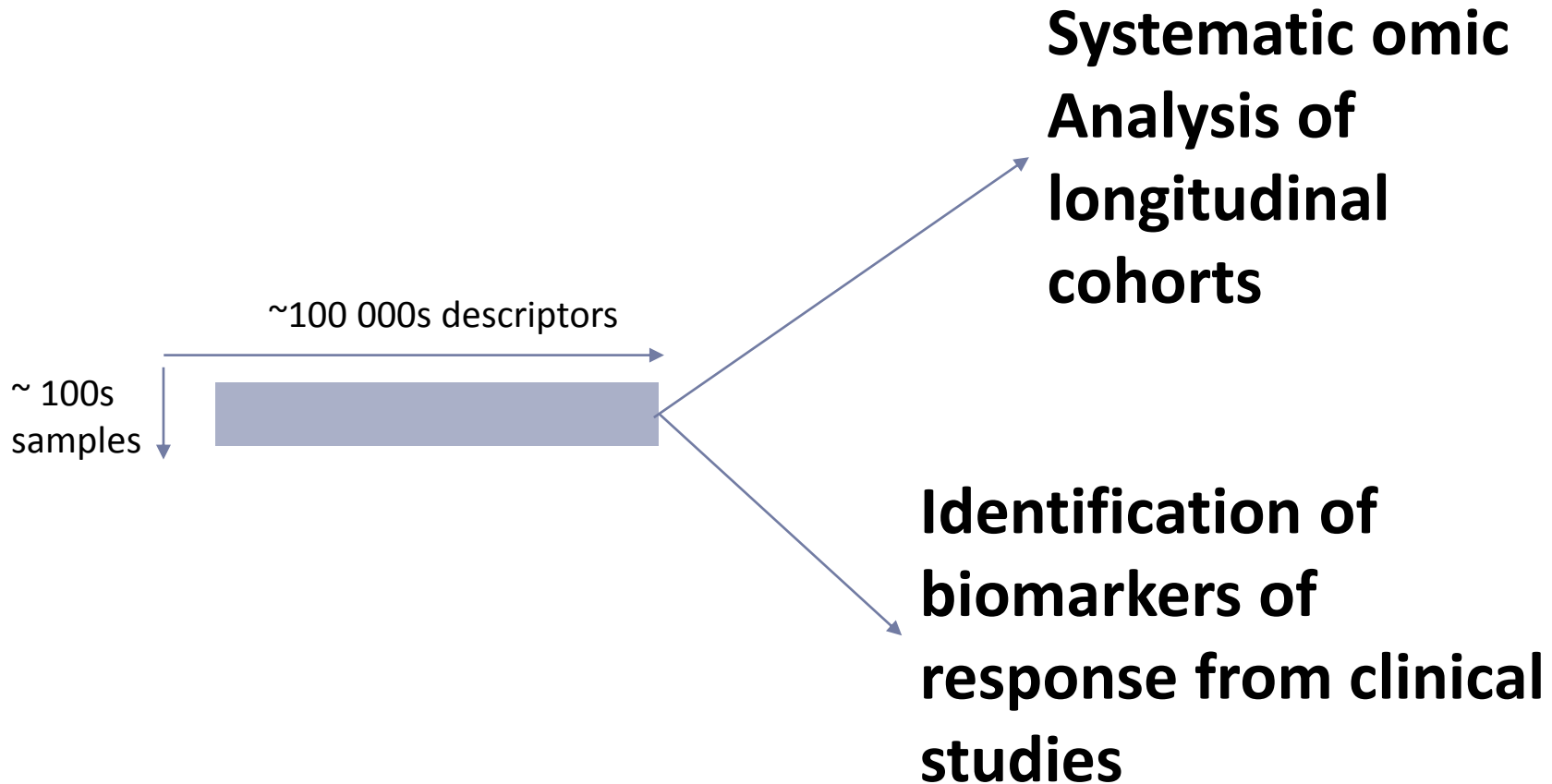
Hypothesis free AI needed

Which AI ? Going beyond the hype…

Association rules: finding strong relations missed by statistics

?

Many things!   N= 1000

N= 1000
P=1%

N= 10
P=90%

Principles of KEM® Galois Lattices or Formal Concept Analysis

Association rules and interestingness measures are key drivers of the approach

KEM®

➢ Unsupervised data mining tool based on Galois Lattices (Formal Concept Analysis)
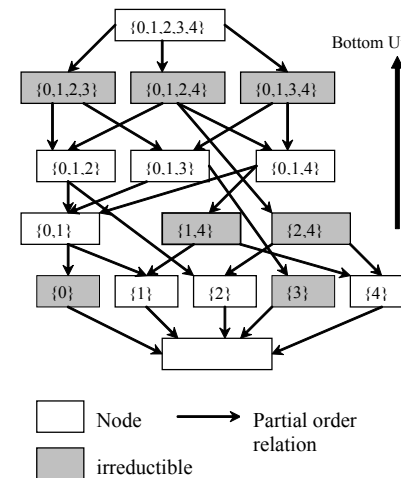
Association rules

➢ Comprehensively analyze complex datasets

➢ Reveal hidden relationships

➢ Derive new hypotheses

Interestingness measures

➢ Measures all logical relations within a dataset

➢ Identifies the most relevant and powerful relations

➢ Decides if and how to use relations to answer a given question

Positioning of KEM® and other data analysis techniques: avoiding over-fitting, data driven (rather than hypothesis) and interpretable
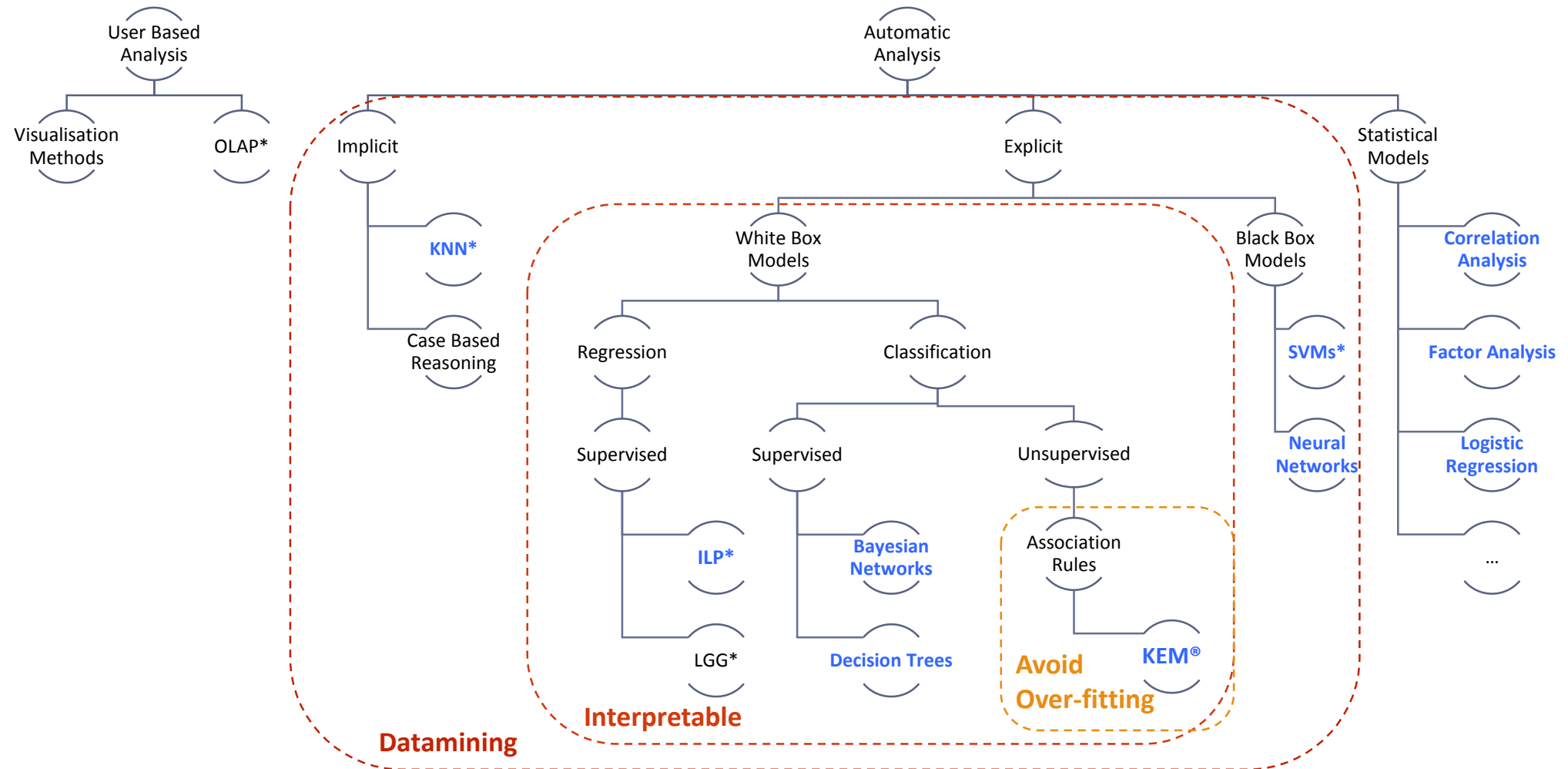
KEM® vs. other data analysis techniques:

1. **Unsupervised** (hypothesis-free) rather than supervised (e.g. optimizing a discrimination objective) to avoid over-fitting;

2. **Interpretable** (white box) rather than black box;

3. **Datamining** rather than statistical (frequentist).

Diagram nodes:
User Based Analysis — Visualisation Methods, OLAP*
Automatic Analysis — Implicit (KNN*, Case Based Reasoning), Explicit, Statistical Models (Correlation Analysis, Factor Analysis, Logistic Regression, …)
White Box Models — Regression (Supervised — ILP*, LGG*), Classification (Supervised — Bayesian Networks, Decision Trees; Unsupervised — Association Rules, KEM®)
Black Box Models — SVMs*, Neural Networks

Datamining — Interpretable — Avoid Over-fitting

* **OLAP** – OnLine Analytic Processing // **KNN** – k-nearest neighbor algorithm // **ILP** – Inductive Logic Programming // **LGG** – Least General Generalizations // **SVMs** – Support Vector Machines

Explainable AI: The next stage of human-machine collaboration

AUGUST 31, 2018

accenture >

RESEARCH REPORT

In brief

- Many artificial intelligence applications today are effectively "black boxes" lacking the ability to "explain" the reasoning behind their decisions.

- As AI expands into areas with large impact on people, such as health care, it will be critical to subject the technology to greater human scrutiny.

- Explainable AI won't replace human workers; rather, it will complement and support people, so they can make better, faster, more accurate decisions.

Data and Analysis plan

Cognitive and memory tests
Evolution of 8 tests monitored
across 6 time points
(Follow-up: 3 years)

For each test, the **delta** of the scores between baseline (M0) and each available time point (**M6**, **M12**, **M18**, **M24**, **M30**, and **M36**) was calculated

A total of 54 delta of scores of cognitive/memory tests was generated

Genotyping: 486,137 variants (SNP) measured,
295,995 present in at least 2 individuals
retained for analysis

Remaining variants are then clustered together:
2 genetic variants shared exactly by the same subjects are grouped together

295,995 variants -> 288,651 clusters

Variant status in subject's genome	Value in data
Absent	0
Heterozygous	1
Homozygous	2

FAB: Frontal Assessment Battery; FCSRT: Free and Cued Selective Reminding Test;
MBT: Memory Binding Test; MMSE: Mini-Mental State Examination; TMT: Trail Making Test

SNP: Single Nucleotide Polymorphism

12

Data and Analysis plan

Neuro Imaging: brain metabolism, volume, connectivity, brain amyloid burden, and cortical thickness

Imaging descriptors were normalized using z-score, and the delta ratio of z-score between baseline (M0) and 24th month (M24) was calculated.

A total of 301 delta ratio of neuroimaging variables were generated

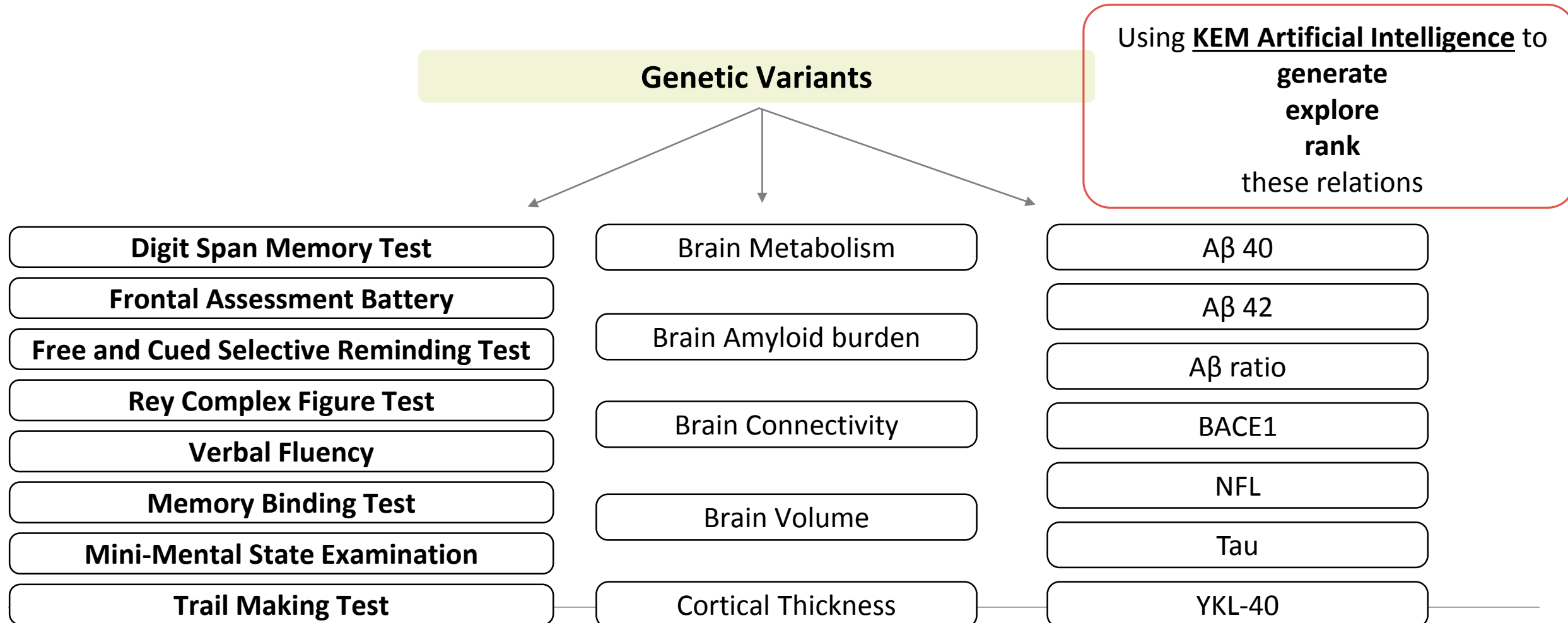Blood based markers: Aβ 40, Aβ 42, Aβ ratio, Bace1, NFL, Tau, YKL-40 at 2 time points

Blood based markers values were normalized using z-score, and the delta ratio of z-score between baseline (M0) and each available time point (M12/M18 and M24) were calculated.

Data and Analysis plan — Linking variants with evolution of scores

The number of relations that can be extracted from the data can be estimated to be 6.6×10^{11}

Within these, 7.4×10^8 are the relations linking genetic variants to the following End Points:

Genetic Variants

Using **KEM Artificial Intelligence** to
generate
explore
rank
these relations

Digit Span Memory Test	Brain Metabolism	Aβ 40
Frontal Assessment Battery		Aβ 42
Free and Cued Selective Reminding Test	Brain Amyloid burden	Aβ ratio
Rey Complex Figure Test		BACE1
Verbal Fluency	Brain Connectivity	NFL
Memory Binding Test		Tau
Mini-Mental State Examination	Brain Volume	
Trail Making Test	Cortical Thickness	YKL-40

Results

1,078 genes (1,158 variants) were selected
Expressed in brain
& Clinically relevant
& Linked to an observed variation on, at least, one imaging biomarker

301 brain imaging descriptors,
54 cognitive scores,
14 blood based biomarkers

Variants
(1158)

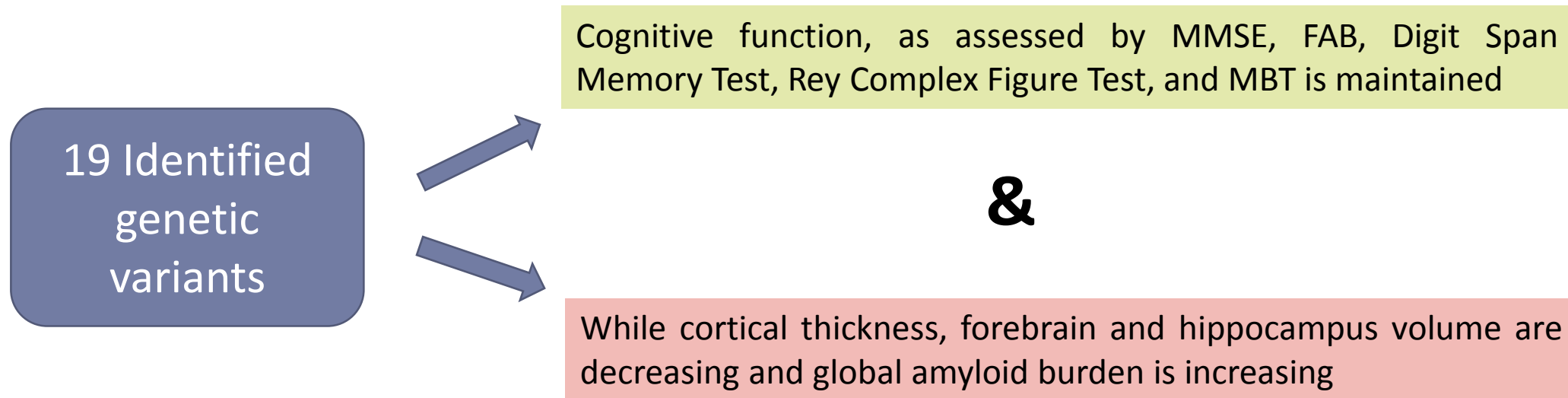|		
Digit Span	Brain Metabolism	Aβ 40
Frontal Assessment Battery		Aβ 42
Free and Cued Selective Reminding Test	Brain Amyloid burden	Aβ ratio
Rey Complex Figure		Bace1
Fluency	Brain Connectivity	NFL
Memory Binding Test	Brain Volume	Tau
Mini Mental State Examination		YKL-40
Trail Making Test	Cortical Thickness	

Variant in rule must hit all 3 categories of endpoints

19 genetic variants
involved in 26 relations

Results

Cognitive function is maintained, while imaging biomarkers show negative evolution

19 Identified genetic variants

→ Cognitive function, as assessed by MMSE, FAB, Digit Span Memory Test, Rey Complex Figure Test, and MBT is maintained

&

→ While cortical thickness, forebrain and hippocampus volume are decreasing and global amyloid burden is increasing

MMSE = Mini-Mental State Examination
FAB = Frontal Assessment Battery
DSMT = Digit Span Memory Test
MBT = Memory Binding Test

Conclusion & Take Home Message

- An complex dataset, consisting of genomic, neuroimaging, cognitive tests and blood based markers, collected in a cohort of subject at risk, was successfully integrated and analyzed
- variants identified as linked **simultaneously** to multiple categories of endpoints: neuroimaging, cognitive tests and blood based markers

- **19 genetic variants** in 253 subjects suggest that compensatory mechanisms at the cell or tissue level brain to maintain cognitive function, despite structural modifications and accumulation of amyloid plaques

- **Artificial Intelligence and Association Rules** are powerful tools to uncover pertinent hypotheses from dataset extensively characterized using molecular and imaging descriptors

- More detailed analysis of **omic features linked to imaging data** will be presented by Dr F. Parmentier) **Thursday** at 11:45   515 AB - O5-07-04

- Application to **Exploring Gut Microbiota as a Source of Potential Biomarkers** (ANAVEX®2-73 Alzheimer's Disease Clinical Study) **Wednesday** 8:45 AM Concourse Hall 152 -   O4-02-04

Thank you

Professor Harald Hampel's Team

Sorbonne University
Department of Neurology
Institute of Memory and Alzheimer's Disease (IM2A)
Brain & Spine Institute (ICM)
François Lhermitte Building
Pitié-Salpêtrière Hospital, Paris, France

Ariana Pharma, Paris, France
Mohammad Afshar
Frédéric Parmentier
Martin Kindermans
Adrien Etcheto
Coralie Williams

www.arianapharma.com

Alzheimer Precision Medicine Initiative
https://www.apmiscience.com

Thank you